拟黑多刺蚁雌激素相关受体基因 cDNA 的克隆及生物信息学分析

欧阳霞辉[1,2,*]，赵有红[3]，侯兰新[1]，奚耕思[2]

(1. 西北民族大学生命科学与工程学院，兰州 730030；2. 陕西师范大学生命科学学院，西安 710062；3. 甘肃省妇幼保健院，兰州 730050)

摘要：雌激素相关受体(estrogen related receptors, ERRs)能够直接与类固醇激素受体共激活子(steroidhormone receptor coactivator, SRC)结合，激活靶基因的表达，与众多生理和发育过程有关。采用 RT-PCR、RACE 等方法从拟黑多刺蚁 *Polyrhachis vicina* Roger 克隆得到了雌激素相关受体基因的 cDNA 克隆，命名为 pvERR (GenBank 登录号为 EF474463)，并采用生物信息学方法对其 cDNA 和编码的蛋白质的理化性质、蛋白质二级及三级结构、分子系统进化关系等进行了预测和推断。结果表明：*pvERR* 基因 cDNA 全长 1 935 bp，包含一个 1 305 bp 的开放阅读框、245 bp 的 5′-UTR 和 385 bp 的 3′-UTR，编码一个由 434 个氨基酸组成的蛋白质。pvERR 的配体结合区(ligand binding domain, LBD)主要由两个 β 折叠和 11 个 α 螺旋(H1, H3 ~ H12, 缺少 H2)构成，这与哺乳类动物已知晶体结构的 ERRγ 的配体结合区结构非常相似。pvERR 氨基酸序列与其他昆虫 ERRs 氨基酸序列同源性很高，它与西方蜜蜂 *Apis mellifera* 雌激素相关受体 amERR 氨基酸序列相似性达到 89.9 %；在进化关系上，pvERR 与人类 *Homo sapiens* ERRs 的关系比果蝇 *Drosophila melanogaster* 雌激素相关受体 dERR 与人类的更近。本文可为进一步研究 pvERR 在昆虫发育中的功能提供有价值的信息。

关键词：拟黑多刺蚁；pvERR；分子克隆；生物信息学分析；蛋白质结构预测

中图分类号：Q966 **文献标识码**：A **文章编号**：0454-6296(2010)05-0587-09

Cloning and bioinformatic analysis of cDNA of estrogen related receptor gene from *Polyrhachis vicina* Roger (Hymenoptera: Formicidae)

OUYANG Xia-Hui[1,2,*], ZHAO You-Hong[3], HOU Lan-Xin[1], XI Geng-Si[2] (1. College of Life Science and Engineering, Northwest University for Nationalities, Lanzhou 730030, China; 2. College of Life Science, Shaanxi Normal University, Xi'an, 710062, China; 3. Gansu Provincial Maternity and Child-care Hospital, Lanzhou, 730050, China)

Abstract: Estrogen related receptor (ERR) can bind to the member of steroid receptor co-activator family and activate transcription activity of the target genes. ERRs have been found to be present in many animal species and play important roles in regulation of many physiological processes. A full-length cDNA of the ERR gene, named *pvERR* (GenBank accession no. EF474463), was cloned from *Polyrhachis vicina* Roger by using RT-PCR and RACE methods. The nucleotide sequence and deduced amino acid sequences of *pvERR* were analyzed and characterized by bioinformatic methods. The results showed that the full-length cDNA of pvERR gene is 1 918 bp, containing an open reading frame of 1 305 bp, 5′-UTR of 245 bp and 3′-UTR of 368 bp. The cDNA encoded a protein with 434 amino acids. The ligand-binding domain (LBD) of pvERRs consists of two β-sheets and eleven α-helixs (H1, H3 – H12, lack of H2) and the LBD structure is similar to that of mammalian ERRγ, whose crystal-structure is known. The homology analysis suggests that the pvERR and the amERR proteins share 89.9% identities in their amino acid sequences, and the homology between them is the highest in all known insect ERRs; phylogenetic relationgship between pvERR and human ERRs is closer than that between *Drosophila* dERR and human ERRs. This study may provide valuable information for further research on pvERR.

Key words: *Polyrhachis vicina*; pvERR; molecular cloning; bioinformatic analysis; protein structure prediction

基金项目：陕西省自然科学基金项目(2008K02-10)

作者简介：欧阳霞辉，女，1978 年生，甘肃靖远人，博士，副教授，主要从事昆虫生殖与发育方面的研究

* 通讯作者 Corresponding author, E-mail: oyxhui316@126.com

收稿日期 Received: 2010-01-02；接受日期 Accepted: 2010-03-29

在细胞发育过程中，各种不同的细胞外信号和它们的配体分子具有重要作用，它们通过控制细胞发育中基因表达，调节细胞的增殖和分化。雌激素以自由扩散形式通过质膜，与核因子受体结合，通过与靶基因上的雌激素应答元件结合来调节基因表达或与其他核蛋白相互作用改变基因的转录活性。雌激素相关受体(estrogen related receptor，ERR)属于孤儿核因子受体，在结构上和雌激素受体(estrogen receptor，ER)非常类似(Pettersson *et al.*, 1996)，它们和 ER 共同构成核因子受体超家族的第 3 亚族，能够直接与类固醇受体共激活子结合激活靶基因的表达(Giguère *et al.*, 1988)。与 ER 不同的是，它们不与雌激素结合，但能够与雌激素应答元件结合(Giguère *et al.*, 1988；Pettersson *et al.*, 1996)，因此，与共同的共激活子作用时，ERR 和 ER 之间存在相互竞争作用(Fujimoto *et al.*, 2005)。而且多项研究表明，ERRs 参与了 ER 信号通路的调节(Bonnelye *et al.*, 2001；Lu *et al.*, 2001；Huss *et al.*, 2002；Horard and Vanacker，2003；Luo *et al.*, 2003；Schreiber *et al.*, 2003；Carrier *et al.*, 2004)，与许多生理和发育过程有关，并参与了细胞的增殖与分化过程(Pettersson *et al.*, 1996；Luo *et al.*, 1997；Mitsunaga *et al.*, 2004；Cheung *et al.*, 2005)。

尽管在人 *Homo sapiens* 和多种动物(小鼠 *Mus musculus*、斑马鱼 *Danio rerio*、文昌鱼 *Branchiostoma floridae* 和海鞘 *Ciona intestinalis* 等)中发现了 ERRs 基因，但在昆虫类，目前仅发现有对黑腹果蝇雌激素相关受体(*Drosophila melanogaster* estrogen related receptor，dERR)的报道(Maglich *et al.*, 2001；Ostberg *et al.*, 2003)。虽然 GenBank 中已收录了冈比亚按蚊 *Anopheles gambiae*、西方蜜蜂 *Apis mellifera*、丽蝇蛹集金小蜂 *Nasonia vitripennis* Walker、赤拟谷盗 *Tribolium castaneum* Herbst 的基因序列和埃及伊蚊 *Aedes aegypti*、斯氏按蚊 *Anopheles stephensi*、库蚊 *Culex pipiens*、拟暗果蝇 *Drosophila pseudoobscura* 的 ERR 蛋白质序列，但是有关昆虫中 ERRs 基因功能的研究未见报道。本研究以具有药用和食用价值的社会性昆虫拟黑多刺蚁 *Polyrhachis vicina* Roger 为研究对象，对其 ERR 基因进行克隆并进行生物信息学分析，以期为该基因的系统演化分析和进一步的研究其在昆虫发育中的功能奠定基础。

1 材料和方法

1.1 实验材料

本研究以拟黑多刺蚁雄蚁为研究对象。拟黑多刺蚁(整窝)购自河北邯郸蚁力王有限公司，在实验室进行饲养。饲养条件：温度 25 ~ 28℃，相对湿度 80%，光周期 12L∶12D。饲料为 20% 的白糖水和豆饼菜叶。

1.2 总 RNA 的提取及 cDNA 第一链的合成

采用 Chmezynski 的酸性异硫氰酸胍：酚：氯仿一步法提取拟黑多刺蚁雄蚁的总 RNA，直接用于 RT-PCR 扩增或 -80℃保存。以 oligo-dT 为引物，参照 MBI 反转录试剂盒 RevertAid™ First-Strand cDNA Synthesis Kit 说明书合成 cDNA 第一链，用于 PCR 扩增或 -20℃保存。

1.3 PvERR cDNA 全长序列获得

由于 ERRs 基因序列接近 2 kb，所以将中间部分序列分成 A，B，C 3 个互相有重叠的片段进行扩增。根据 GenBank 中已知的昆虫 ERRs 基因序列和蛋白质序列，在 ERRs 保守的 DNA 结合区(DNA binding domain，DBD)和配体结合区(ligand binding domain，LBD)设计了一对兼并引物，AE1a (5′-CAAGGHAAYATCGAGTACA-3′) 和 AE1b (5′-ATCTGYTTCGCCCAVCCRAT-3′)，扩增片段 A。将 A 片段测序以后通过在 NCBI 上 Blast 确定是 ERR 的同源序列后，再根据 A 片段设计 2 个特异性引物 AE2b (5′-CCAGCCGATTATGCCAACCAA-3′) 和 AE3b (5′-CATYTCDACRAASAGCTTGTTCAT-3′)，分别与根据保守区设计的兼并引物 AE2a (5′-AGGTAKCCCTSGAASTCC-3′)和 AE3a (5′-CATYTCDACRAASAGCTTGTTCAT-3′) 组合扩增片段 B 和 C，然后将 3 个片段拼接获得大部分编码序列。

采用 TaKaRa RACE 试剂盒 (3′-Full RACE Core Set and 5′-Full RACE Core Set) 扩增 PvERR 基因两端序列，操作完全按照说明书进行。在 5′ RACE 实验中，按照 RACE 引物设计原则在已知序列中设计了一条 5′磷酸化的特异性反转录引物和两对特异性的巢式 PCR 引物 (RT-5P: 5′-TCCTTCAGCATTCCT-3′; 5AE1a: 5′-GTGAGATAAATAAACGCAAACG-3′; 5AE1b: 5′-ACGCTTCGCACGAGGCAAC-3′; 5AE2a: 5′-CAAGGCATGTAGATTTC

AGA-3′；5AE2b：5′-AGACAGAGTCTTCTGGGTATCA-3′）；在 3′RACE 实验中，按照实验原理在已知序列中设计了两个特异性引物，分别与两个 3′通用引物组合成对(3AE1：5′-CCACGCAGAATATGTTCCTA-3′；3AE2：5′-ATTGGTTGGGCGAAACAGAT-3′）。本研究中所有产物片段的测序均由上海生工生物工程有限公司完成。

1.4　序列分析

通过 NCBI 中在线分析工具 ORF Finder（http：//www. ncbi. nlm. nih. gov/gorf/gorf. html）和 PSI and PHI-BLAST（Position-specific iterated and pattern-hit initiated BLAST）（http：//blast. ncbi. nlm. nih. gov/Blast. cgi）进行开放阅读框查找和同源性分析（Altschul *et al.*，1997；Schäffer *et al.*，2001）；蛋白质的理化性质和疏水性分析采用在线软件 Protparam（http：//www. expasy. ch/tools/protparam. html）和 Protscal（http：//www. expasy. ch/tools/protscale. html）。蛋白质二级结构的预测采用 APSSP2（Advanced Protein Secondary Structure Prediction Ser-ver）（http：//imtech. res. in/raghava/apssp/）网上软件在线预测（Raghava，2000）；蛋白质三级结构预测采用 Swiss Model First Approach Mode（http：//swissmodel. expasy. org/）在线软件预测（Schwede *et al.*，2003）。

1.5　系统进化分析

以低等脊椎动物文昌鱼 *Branchiostoma Floridae*（AAU88062. 1）和现代人 *H. sapiens*（NP_004443. 2）的 ERR 作为外群，对已经提交 GenBank 的 9 种昆虫 ERRs 氨基酸序列（黑腹果蝇 *D. melanogaster*，NP_729340，西方蜜蜂 *A. mellifera*，XP_392385，斯氏按蚊 *A. stephensi*，BAE96770，拟暗果蝇 *D. pseudoobscura*，XP_001354210，冈比亚按蚊 *A. gambiae*，XP_321343，丽蝇蛹集金小蜂 *N. vitripennis*，XP_001604033，赤拟谷盗 *T. castaneum*，XP_001812322，库蚊 *C. pipiens*，XP_001862599，埃及伊蚊 *A. aegypti*，EAT34188）和 PvERR 氨基酸序列进行系统进化分析。通过 CLUSTALX 对序列进行比对，然后采用 MEGA（Ver 4. 1）软件基于 NJ（neighbor-joining）法构建进化树，距离校正使用泊松模型，计算时将空位完全删除，自举分析通过1 000次循环实现。

2　结果与分析

2.1　拟黑多刺蚁 PvERR 的 cDNA 序列、氨基酸组成和理化性质分析预测

2.1.1　cDNA 序列分析和预测蛋白质的氨基酸组成：拟黑多刺蚁雌激素相关受体基因 cDNA 全长 1 935 bp，包含一个 1 305 bp 的开放阅读框、245 bp 的 5′-UTR 和 385 bp 的 3′-UTR，编码一个由 434 个氨基酸组成的蛋白质（图 1）。5′-UTR 有一个框内终止密码子，3′-UTR 有一个 17 bp 的 ploy（A）尾巴。编码的蛋白在 5′端有一个起始密码子 ATG，在 3′端有一个终止密码子 TGA。在 ATG 密码子附近发现最常见的翻释起始共有序列：5′-GCCATGG-3′，即 Kozak 序列。将拟黑多刺蚁雌激素相关受体基因命名为 *pvERR*（GenBank 登录号为 EF474463）。

2.1.2　理化性质：编码蛋白质 pvERR 分子量为 49. 0989 kD，理论等电点为 6. 62。氨基酸组成中含强酸氨基酸（Asp，Glu）53 个，强碱氨基酸（Arg，Lys）52 个，极性氨基酸 247 个，疏水氨基酸 187 个。pvERR 氨基酸组成中 Leu 含量最高，这与其他 ERRs 蛋白是一致的；与人的 ERRs 相比，pvERR 的氨基酸组成中 Leu，Lys，His 和 Ser 所占的百分比低于人 ERRs，而 Gln，Thr，Cys 和 Trp 所占的百分比高于人 ERRs。

2.1.3　序列同源性与多序列比对分析：将 pvERR 氨基酸序列在 NCBI 上进行 PSI and PHI-BLAST 分析，对非冗余数据库的搜索显示，此蛋白序列与数据库中其他动物的 ERRs 蛋白质序列同源性很高，例如，它与西方蜜蜂的 amERR 氨基酸序列一致性为 89. 9%；它与果蝇、埃及伊蚊和斯氏按蚊的 ERR 氨基酸序列一致性分别为 50. 2%，50. 7% 和 52. 1%；与人的 ERRα 的氨基酸序列一致性也达到了 41. 8%。

对拟黑多刺蚁 pvERR 蛋白质序列进行 Conserved Domain（CD）搜索（Marchler-Bauer and Bryant，2004），预测到两个保守的功能结构域，即核受体 C4 锌指结构（ZnF_C4）和核受体配体结合结构域（HOLI）。在 pvERR 蛋白质中，DNA 结合区 ZnF_C4 是由 73 个氨基酸编码（105 ~ 177），配体结合区的 HOLI 由 158 个氨基酸编码（244 ~ 401），这与西方蜜蜂 amERR（XP_392385）、果蝇 dERR（NP_729340）

```
1     atgctgaagagtcgcgtacggcgcgcggtgccgtcgacgtcgtctgtgttggcagcggtggt
63    agtgggccctgtatccgcgcgagcggcgcgctcgcccgcgtcgaggggaatcccatccgg
123   aggacccatccgaggatctcttggagtagattcgcgagccgagtatcgcgtcctgccacg
183   actgtcaaatagtcgggcaggtgacaggaggtgatttctgtcggggacgaggagacgag
243   gccatggacgcgtggatgtacgatgtggtttgtatgatgtctgacgctaccacggaaagc
        M  D  A  W  M  Y  D  V  V  C  M  M  S  D  A  T  T  E  S
303   atgatcgggaataacaggacgatgtccaacatcaagcaggaaatcgataacccaacgacg
      M  I  G  N  N  R  T  M  S  N  I  K  Q  E  I  D  N  P  T  T
363   cctacgcaaaattatcaagtatgctcgccgaccacgactctccagcatcaggagtcgata
      P  T  Q  N  Y  Q  V  C  S  P  T  T  T  L  Q  H  Q  E  S  I
423   tgcacgaaattggatattccgcaggattatggcggtagcaacggtagccctgaaagccca
      C  T  K  L  D  I  P  Q  D  Y  G  G  S  N  G  S  P  E  S  P
483   gagatgcatcattgttcctcgacgacgcagcctttaggtactccagaggatggtattaaa
      E  M  H  H  C  S  S  T  T  Q  P  L  G  T  P  E  D  G  I  K
543   gaagaggatatgatacccagaagactctgtctagtttgcggtgatgtcgcgagcggtttt
      E  E  D  M  I  P  R  R  L  C  L  V  C  G  D  V  A  S  G  F
603   cactacggggttgcctcgtgcgaagcgtgcaaggctttcttcaagagaactatacaagga
      H  Y  G  V  A  S  C  E  A  C  K  A  F  F  K  R  T  I  Q  G
663   aacatcgagtatacttgtccggcgaatggcgaatgtgagataaataaacgcaaacggaag
      N  I  E  Y  T  C  P  A  N  G  E  C  E  I  N  K  R  K  R  K
723   gcgtgccaggcatgtagatttcagaaatgtttgcggcaaggaatgctgaaggaaggcgtg
      A  C  Q  A  C  R  F  Q  K  C  L  R  Q  G  M  L  K  E  G  V
783   cgattggatcgcgttaggggcggtcgacaaaagtatagaagatccgcggatccttacacg
      R  L  D  R  V  R  G  G  R  Q  K  Y  R  R  S  A  D  P  Y  T
843   ccggtcaagagcgccacgcttgaagataacaagttagtagaagccttagcagcttgcgag
      P  V  K  S  A  T  L  E  D  N  K  L  V  E  A  L  A  A  C  E
903   ccagatatgcttcaagtatccaacatttccagcacacaggaaacagaccaaaaagttctc
      P  D  M  L  Q  V  S  N  I  S  S  T  Q  E  T  D  Q  K  V  L
963   ggacaattgtcggatctttacgaccgagaattggttggcataattggttgggcgaaacag
      G  Q  L  S  D  L  Y  D  R  E  L  V  G  I  I  G  W  A  K  Q
1023  atcccaggattcagcagcttagcgttaaatgatcaaatgcgacttttgcagagtacttgg
      I  P  G  F  S  S  L  A  L  N  D  Q  M  R  L  L  Q  S  T  W
1083  gctgagatattaacttttactctagcgtggagaagcacaccaaataatggcagattgagg
      A  E  I  L  T  F  T  L  A  W  R  S  T  P  N  N  G  R  L  R
1143  tttgcgcaagattttacattagatgaaaggattgcgcgcgattgtcactgtatggaactt
      F  A  Q  D  F  T  L  D  E  R  I  A  R  D  C  H  C  M  E  L
1203  tacacacattgtattcaagtcgtcgagaggatccagcgattgggtttaacgcgggaggaa
      Y  T  H  C  I  Q  V  V  E  R  I  Q  R  L  G  L  T  R  E  E
1263  tattacgtgttgaaagcattgatcttagcgaacagcgatgctagatctgacgaaccacaa
      Y  Y  V  L  K  A  L  I  L  A  N  S  D  A  R  S  D  E  P  Q
1323  gcactttaccgttttcgtgattctatcttgaattccctgtcggactgtgtggcggcagtg
      A  L  Y  R  F  R  D  S  I  L  N  S  L  S  D  C  V  A  A  V
1383  agacccggtcaggctcttcgtgccacgcagaatatgttcctagtgttaccaagtcttaga
      R  P  G  Q  A  L  R  A  T  Q  N  M  F  L  V  L  P  S  L  R
1443  caggccgacggaattgtcagacgattttggtctagtgtgtacagaacgggcaaggtaccg
      Q  A  D  G  I  V  R  R  F  W  S  S  V  Y  R  T  G  K  V  P
1503  atgaataagttatttgaggagatgctcgagaacgtttgtaaccgatgaaacatgcgttgt
      M  N  K  L  F  E  E  M  L  E  N  V  C  N  R  *
1563  aaatttgataattcgactattcgacacctgttcgtgcattgtttatcgaaaattggcaat
1623  gttcttcatcattaatgtgatcattattatcattgtacatcgttatcgaatacgaaatcg
1683  ttttggtgtcgataagtgatagtgcgtgagaaaaatgaagttatcgatttttttacgatgc
1743  tttttattttctaatggactcgtgaatctttatttttttcttttggaaagaatgttgaaa
1803  gtgaaaatttggctaatgtctccggcggtaatcgcgaatcgcgaatcgttccagaacggt
1863  aggcgcgtctttgtttgacgtatcctttccttttcgagtaccccatttacaggtacaaaa
1923  aaaaaaaaaaaaa
```

图 1　拟黑多刺蚁 *pvERR* 基因核苷酸序列及其所编码的蛋白的氨基酸序列

Fig. 1　Nucleotide and deduced amino acid sequence of *ERR* from *Polyrhachis vicina*

矩形框内的序列为 Kozak 序列，下划线标出的是框内终止密码子。Rectangular frame indicates the Kozak sequence, and the underline indicates the in-frame stop codon.

是完全一样的（图 2）。通过比对发现，pvERR，amERR，dERR 的 DBD 和 LBD 分别和人 ERR 的对应部分相似性都达到 85% 和 40% 以上，可见 ERR 的这两个功能区在演化上是非常保守的。

2.1.4　二级结构预测：采用 APSSP2 法对 pvERR 蛋白质的二级结构进行了网上在线预测，结果如图 3 所示。和多数核受体一样，pvERR 分子结构由位于 N 端的 A/B 区域，DNA 结合域（DBD）、铰链区 D 和配体结合域（LBD）组成。pvERR 蛋白的 DNA 结合区（105 ~ 177），主要由 4 个 β 折叠和 3 个 α 螺旋构成。在 pvERR 蛋白的配体结合区（206 ~ 429），可以明显看到两个 β 折叠和 11 个 α 螺旋（H1 和 H3 ~ H12）。

2.1.5　三级结构预测：采用 SwissModel First Approach Mode 在线基于同源建模软件对 pvERR 蛋白质的三级结构进行了预测。由于到目前为止，结构数据库 PDB 中没有任何一个物种的完整全长 ERRs 蛋白质的结构模型，所以以已知的 ERR DBD 和 LBD 结构模型，对拟黑多刺蚁 pvERR 蛋白质的 DBD 和 LBD 结构分别进行了三级结构预测。

预测 pvERR DBD 区采用的模版是人 ERRβ DBD（PDB ID：1LO1，NMR Resolution：99.9），序列（104 ~ 192）一致性（Sequence Identity）为84%，评价值为 1.17e-41，预测结果见图 4。pvERR DBD 区由于与模板的序列匹配度达到 84%，所以预测的结构与模板差异很小，属于 α + β 型，都有 4 个 β 折叠和 3 个 α 螺旋，由位于 109，112，126，129 位的半胱氨酸（Cys）形成的第一个锌指结构（ZnF_C4）与靶基因 DNA 的反应元件结合（图 4）。预测 pvERR LBD 采用的模版是人 ERRγ LBD（PDB ID：2gpuA，X-RAY Resolution：1.70），序列（209 ~ 429）一致性（Sequence Identity）为 39%，评价值为 1.77e-37，预测结果见图 5。pvERR LBD 区主要包含 2 个 β 折叠和 11 个 α 螺旋（缺少 H2），由 H3，H4 和 H12 形成配体结合袋（图 5）。

```
Polyrhachis_vicina_roger        ----MDAWMYDVVCMMSDATTESMIGNNRTMSNIKQEIDNPTTPTQNYQV
Apis_mellifera                  ----MDSWMYDVVCMMPGGGTENMIGNNRTMANIKQEIENPTTPTQNYQV
Drosophila_melanogaster         MSDGVSILHIKQEVDTPSASCFSPSSKSTATQSGTNGLKSSPSVSPERQL
                                    :.    .     ...   .  .:. :  . .: :....: : *:

Polyrhachis_vicina_roger        CSPTTTLQHQESICTKLDIPQDYGGS--NGSPESPEMHHCSSTTQPLGTP
Apis_mellifera                  CSPTTTLQHQEVICSKIEVPPDYGGG--EGSPGSPEMHHCSSTTQPLGTS
Drosophila_melanogaster         CSSTTSLSCDLHNVSLSNDGDSLKGSGTSGGNGGGGGGTSGGNATNASA
                                **.**:*. :    :  :   .  *.  .*.  .      *. . . .:.

                                                  ZnF_C4
Polyrhachis_vicina_roger        EDGIKEEDMIPRRLCLVCGDVASGFHYGVASCEACKAFFKRTIQGNIEYT
Apis_mellifera                  EEGVKEEDMIPRRLCLVCGDVASGFHYGVASCEACKAFFKRTIQGNIEYT
Drosophila_melanogaster         GAGSGSVRDELRRLCLVCGDVASGFHYGVASCEACKAFFKRTIQGNIEYT
                                  *  .     ***************************************

Polyrhachis_vicina_roger        CPANGECEINKRKRKACQACRFQKCLRQGMLKEGVRLDRVRGGRQKYRRS
Apis_mellifera                  CPANGECEINKRRRKACQACRFQKCLRQGMLKEGVRLDRVRGGRQKYRRS
Drosophila_melanogaster         CPANNECEINKRRRKACQACRFQKCLLMGMLKEGVRLDRVRGGRQKYRRN
                                ****.*******:*************  *********************.

Polyrhachis_vicina_roger        ADPYTPVKSATLED------------NKLVEALAACEPDMLQVSN-----
Apis_mellifera                  TDPYTPVKPAPLEANIATGASNEIINNKMLEALAACEPDMLQVSN-----
Drosophila_melanogaster         PVSNSYQTMQLLYQSN----TTSLCDVKILEVLNSYEPDALSVQTPPPQV
                                . . :  .   *               *::*.* : *** *.*..

Polyrhachis_vicina_roger        --------------------------------ISSTQETDQKVLGQLSDL
Apis_mellifera                  --------------------------------ISHTLDTDQRVLGQLSDL
Drosophila_melanogaster         HTTSITNDEASSSSGSIKLESSVVTPNGTCIFQNNNNNDPNEILSVLSDI
                                                                 . . :  :.:*. ***:

Polyrhachis_vicina_roger        YDRELVGIIGWAKQIPGFSSLALNDQMRLLQSTWAEILTFTLAWRSTPNN
Apis_mellifera                  YDRELVGIIGWAKQIPGFSSLALNDQMRLLQSTWAEILTFSLAWRSMPNN
Drosophila_melanogaster         YDKELVSVIGWAKQIPGFIDLPLNDQMKLLQVSWAEILTLQLTFRSLPFN
                                **:***.:********** .*.*****:*** :******: *::** * *

                                              HOLI
Polyrhachis_vicina_roger        GRLRFAQDFTLDERIARDCHCMELYTHCIQVVERIQRLGLTREEYYVLKA
Apis_mellifera                  GRLRFAQDFTLDERLARECHCTELYTHCIQIVERLQRLGLTREEYYVLKA
Drosophila_melanogaster         GKLCFATDVWMDEHLAKECGYTEFYYHCVQIAQRMERISPRREEYYLLKA
                                *:* ** *. :**::*::*   *:* **:*:.:*::*:.  *****:***

Polyrhachis_vicina_roger        LILANSDARSDEPQALYRFRDSILNSLSDCVAAVRPGQALRATQNMFLVL
Apis_mellifera                  LILANSDARSDEPQALYRFRDSILNSLSDCMAAVRPGQALRATQNMFLVL
Drosophila_melanogaster         LLLANCDILLDDQSSLRAFRDTILNSLNDVVYLLRHSSAVSHQQQLLLLL
                                *:***.*   *: .:*  ***:*****.* :  :* ..*:   *:::*:*

Polyrhachis_vicina_roger        PSLRQADGIVRRFWSSVYRTGKVPMNKLFEEMLENVCNR
Apis_mellifera                  PSLRQVDGIVRRFWSSVYRTGKVPMNKLFVEMLEAAYYR
Drosophila_melanogaster         PSLRQADDILRRFWRGIARDEVITMKKLFLEMLEPLAR-
                                *****.*.*:**** .: *   :.*:*** ****
```

图 2　采用 CLUSTAL X 对拟黑多刺蚁、西方蜜蜂和黑腹果蝇雌激素相关受体氨基酸序列的比对结果

Fig. 2　Alignment of amino acid sequences of pvERR, amERR and dERR using CLUSTAL X method

＊表示相同的氨基酸；: 和 . 表示氨基酸的保守性程度；大方括号内的序列分别是 ERRs 的保守结构域 ZnF_C4 和 HOL1；矩形框表示 pvERR 和 dERR 铰链区的差异。“＊” shows the same amino acids; “:” and “.” show the degree of conservation of different amino acids. Sequences in the square bracket show the conserved domain, ZnF_C4 and HOL1 of ERRs. The rectangular frame shows the difference in the hinge region between pvERR and dERR. PvERR: 拟黑多刺蚁雌激素相关受体 ERR from *Polyrhachis vicina* roger; amERR: 西方蜜蜂雌激素相关受体 ERR from *Apis mellifera*; dERR: 黑腹果蝇雌激素相关受体 ERR from *Drosophila melanogaster*.

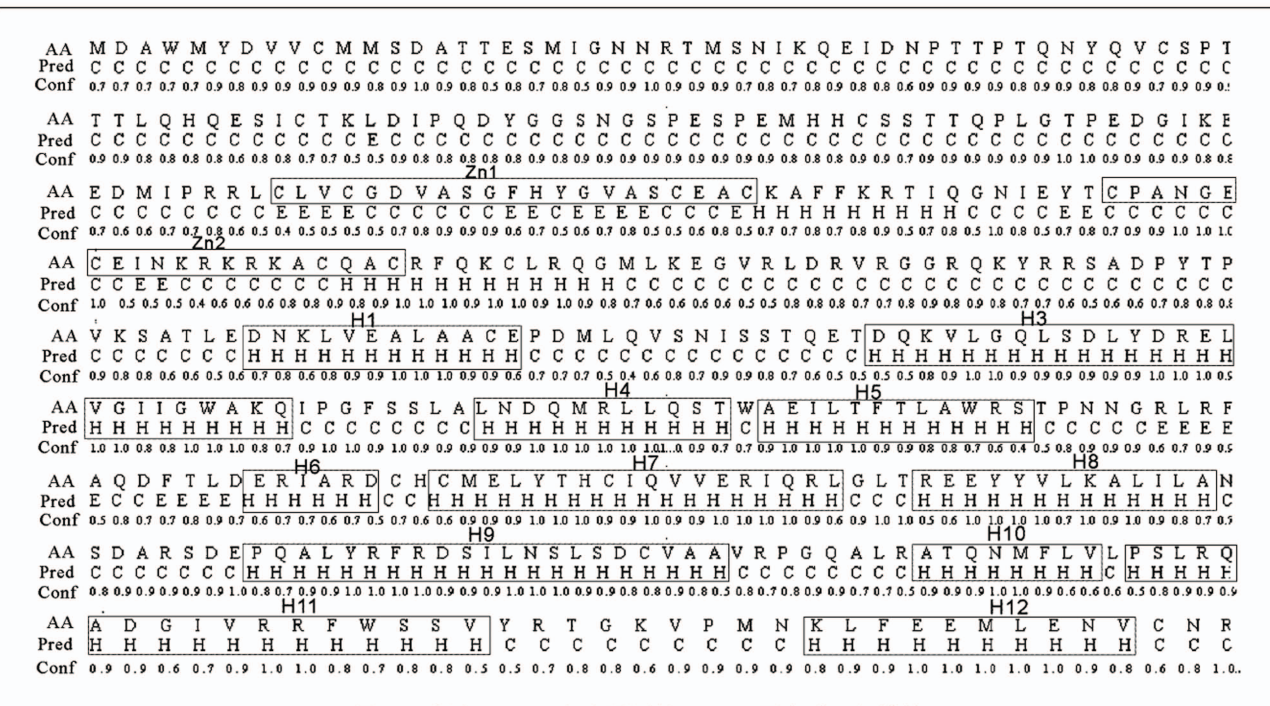


图 3 采用 APSSP2 方法预测的 pvERR 蛋白质二级结构

Fig. 3 Predicted secondary structure of the pvERR protein using APSSP2

AA 表示氨基酸，Pred 表示预测的氨基酸所对应的二级结构状态(H：螺旋；E：折叠 C：卷曲)，Conf 表示预测的准确性。AA represents amino acid, and Pred (Prediction) represents predicted corresponding secondary structure state (H: Helix; E: Strand; C: Coil), and Conf (Confidence) stands for probability of correct prediction.

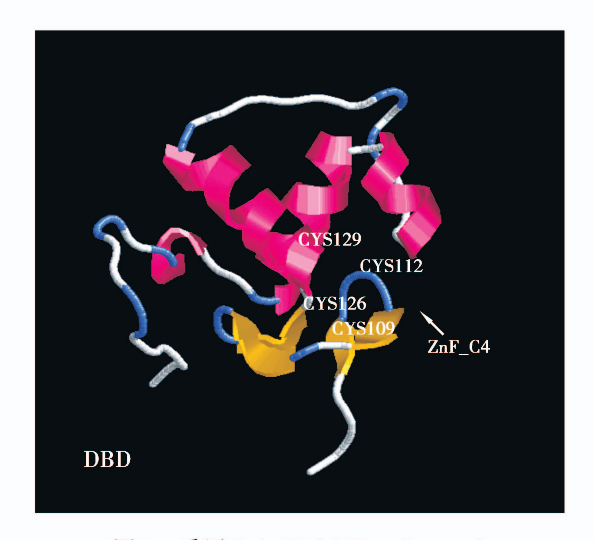


图 4 采用 SwissModel First Approach 预测的 pvERR 蛋白 DBD 区的三维结构

Fig. 4 Predicted tertiary structure of DBD domain in pvERR protein using SwissModel First Approach



图 5 采用 SwissModel First Approach 预测的 pvERR 蛋白 LBD 区的三维结构

Fig. 5 Predicted tertiary structure of LBD domain in the pvERR protein using SwissModel First Approach

2.2 亲水性与疏水性分析

虽然 pvERR 蛋白质整体都具有较强的亲水性，但它的功能区存在多个疏水区，DNA 结合区第一个锌指结构区(RRLCLVCGDVASGFHYGVASCEACK，106～130)是一个明显的疏水区，疏水性最强的位点是 112 位的半胱氨酸(2.1)；配体结合区 LBD (212～397)有 9 个明显的疏水区(图 6)。

2.3 系统发育分析

采用 MEGA(Ver 4.1)软件基于 NJ 法构建了进化树(图 7)。从进化树可以看出，pvERR 和 amERR 同源性很高，它们与所有已知的昆虫 ERRs 聚在了一起；pvERR 和 amERR 与人 ERRs 的关系

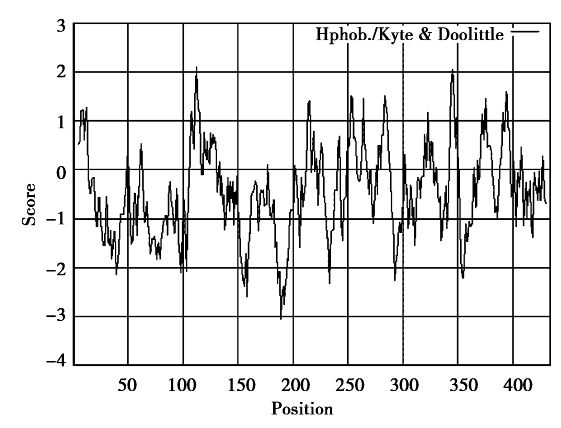


图 6　采用 ProtScale 对 pvERR 蛋白质的疏水性/亲水性预测结果

Fig. 6　Predicted hydrophobicity/ hydrophilicity of pvERR using ProtScale

比双翅目昆虫与人类的更近，也就是说，单从 ERRs 来看，在进化关系上，pvERR 和 amERR 比双翅目昆虫的 ERRs 都要保守。

3　讨论

哺乳动物 ERRs 参与 ER 信号通路的调控，与许多生理和发育过程有关，并参与细胞的增殖与分化（Pettersson *et al.*, 1996；Luo *et al.*, 1997；Mitsunaga *et al.*, 2004 ；Cheung *et al.*, 2005）。近年来的研究证实在昆虫中也存在 ERR（Maglich *et al.*, 2001；Ostberg *et al.*, 2003），但是昆虫中没有雌激素，那么昆虫中 ERR 又是如何发挥作用的呢？本研究对社会性昆虫拟黑多刺蚁的 ERR 进行了克隆并采用生物信息学方法对其全长 cDNA 序列和其编码的蛋白序列进行了分析，为进一步对昆虫 ERR 的系统发生及功能分析提供依据。

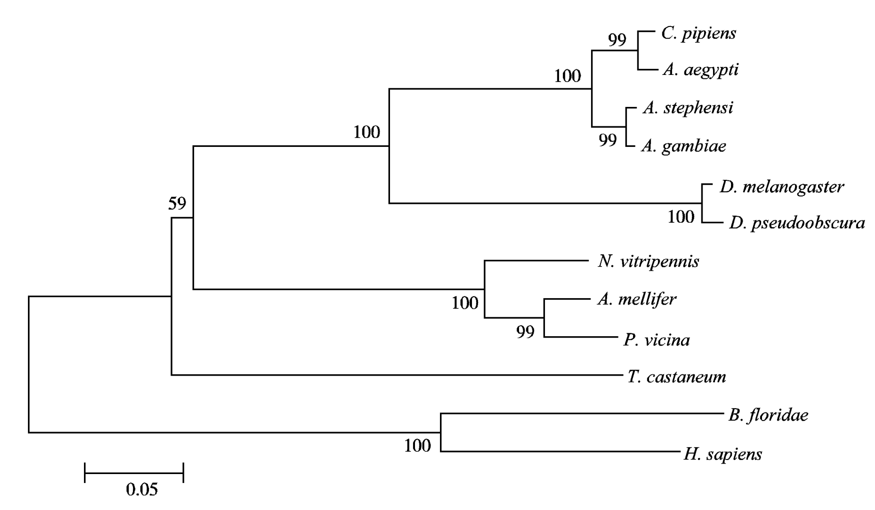


图 7　采用 NJ 法构建的 ERR 系统发育树

Fig. 7　Phylogenetic tree of ERRs based on NJ method

B. floridae：文昌鱼 *Branchiostoma floridae*（AAU88062.1）；*H. sapiens*：人 *Homo sapiens*（NP_004443.2）；*D. melanogaster*：黑腹果蝇 *Drosophila melanogaster*（NP_729340）；*A. mellifera*：西方蜜蜂 *Apis mellifera*（XP_392385）；*A. stephensi*：斯氏按蚊 *Anopheles stephensi*（BAE96770）；*D. pseudoobscura*：拟暗果蝇 *Drosophila pseudoobscura*（XP_001354210）；*A. gambiae*：冈比亚按蚊 *Anopheles gambiae*（XP_321343）；*N. vitripennis*：丽蝇蛹集金小蜂 *Nasonia vitripennis*（XP_001604033）；*T. castaneum*：赤拟谷盗 *Tribolium castaneum*（XP_001812322）；*C. pipiens*：库蚊 *Culex pipiens*（XP_001862599）；*A. aegypti*：埃及伊蚊 *Aedes aegypti*（EAT34188）；*P. vicina*：拟黑多刺蚁 *Polyrhachis vicina*. 节点处数字表示重复 1 000 次的 Bootstrap 值 Bootstrap values by 1 000 repeats were shown in the nodes.

CD 搜索显示拟黑多刺蚁 pvERR 氨基酸序列包含 ZnF_C4 和 HOLI 两个功能结构域，而位于 DNA 结合区的 ZnF_C4 和位于配体结合区的 HOLI 是核激素受体家族所特有的两个保守结构域，所以 pvERR 同其他物种 ERRs 一样属于核因子受体超家族第 3 亚族。拟黑多刺蚁 pvERR 氨基酸序列与数据库中其他动物的 ERRs 氨基酸序列同源性很高，尤其是 DBD 区和 LBD 区，差异主要体现在这两个区之间的铰链区，与果蝇相比，pvERR 与西方蜜蜂的 amERR 铰链区较短（图 2），这可能与物种本身的进化有关。

系统发育分析显示，在进化关系上，pvERR 和 amERR 比双翅目昆虫的 ERRs 都要保守，这样的结果不仅与物种本身的进化关系相符，而且与对蜜蜂基因组一些基因的研究结果是一致的。即尽管相对于其他全基因组已测序的昆虫，蜜蜂基因组进化缓慢、极其古老，但它体内控制生理节奏、衰

老、RNA(核糖核酸)干扰的基因，与果蝇等昆虫的同类基因相差较大，反而与人类等脊椎动物的同类基因更为接近(Honeybee Genome Sequencing Consortium，2006)。

拟黑多刺蚁 pvERR 氨基酸序列二级和三级结构的预测结果显示，疏水性分析中预测到的 DNA 结合区的疏水性区域主要是由 4 个 β 折叠形成的第一个锌指结构(图 4)，该结构主要与靶基因 DNA 的反应元件结合(Giguère，1999)。通常 ERR 的配体结合区是一个由 12 个 α 螺旋(H1 ~ H12)形成的三明治式的结构，通过比对，发现 pvERR 缺少 H2，这与哺乳类已知晶体结构的 ERRγ 的配体结合区结构非常类似(Greschik *et al.*，2002)。

研究表明，ERR 的疏水性配体通常结合于一个称为配体结合袋(ligand binding pocket，LBP)的结构中，配体依赖的转录激活由核受体中位于 LBD 的 H12 的 AF-2 介导，即由 H3，H4 和 H12 形成的疏水表面与共激活子家族的 LXXLL 基序之间的相互作用完成(Feng *et al.*，1998)。果蝇 dERR 中 Y295，T333 和 Y365 3 个位点几乎完全填满了配体结合袋，所以 dERR 活性不能被拮抗剂 4-OHT 它莫西芬(tamoxifen)和 DES 抑制。如果使黑腹果蝇 dERR LBD 中的 3 个位点发生突变，即 Y295A/T333I/Y365L，突变体的 dERR 活性就可以明显被拮抗剂 4-OHT 它莫西芬和 DES 抑制，而且在转染实验中，共激活子 Mgrip-1 (mouse giucocoricoid receptor interacting protein 1)能够显著加强突变体的活性(Ostberg *et al.*，2003)，说明这 3 个位点在与配体结合中具有重要的作用。在 pvERR 氨基酸序列中，由 H3，H4 和 H12 形成 LBP，与果蝇 dERR 中 Y295，T333 和 Y365 3 个位点相对应的位点分别是 Y246，T284 和 C316 (图5)。有关 pvERR 与拮抗剂 4-OHT 它莫西芬和 DES 的关系有待进一步的研究。

参 考 文 献 (References)

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25: 3 389 – 3 402.

Bonnelye E, Merdad L, Kung V, Aubin JE, 2001. The orphan nuclear estrogen receptor-related receptor α is expressed throughout osteoblast differentiation and regulates bone formation *in vitro*. *Journal of Cell Biology*, 153(5): 971 – 984.

Carrier JC, Deblois G, Champigny C, Levy E, Giguère V, 2004. Estrogen related-receptor α (ERRα) is a transcriptional regulator of apolipoprotein A-Ⅳ and controls lipid handling in the intestine. *Journal of Biological Chemistry*, 279(50): 52 052 – 52 058.

Cheung CP, Yu S, Wong KB, Chan LW, Lai FM, Wang X, Suetsugi M, Chen S, Chan FL, 2005. Expression and functional study of estrogen receptor-related receptors in human prostatic cells and tissues. *J. Clin. Endocrinol. Metab.*, 90(3): 1 830 – 1 844.

Feng W, Ribeiro RC, Wanger RL, Nguyen H, Apriletti JW, Fletterick RJ, Baxter JD, Kushner PJ, West BL, 1998. Hormone-dependent coactivator binding to a hydrophobic cleft on nuclear receptors. *Science*, 280(5 370): 1 747 – 1 749.

Fujimoto J, Nakagawa Y, Toyoki H, Sakaguchi H, Sato E, Tamaya T, 2005. Estrogen-related receptor expression in placenta throughout gestation. *Journal of Steroid Biochemistry and Molecular Biology*, 94 (1 – 3): 67 – 79.

Giguère V, 1999. Orphan nuclear receptors: From gene to function. *Endocrine Reviews*, 20(5): 689 – 725.

Giguère V, Yang N, Segui P, Evans RM, 1988. Identification of a new class of steroid hormone receptors. *Nature*, 331(6 151): 91 – 94.

Greschik H, Wurtz JM, Sanglier S, Bourguet W, van Dorsselaer A, Moras D, Renaud JP, 2002. Structural and functional evidence for ligand-independent transcriptional activation by the estrogen-related receptor 3. *Molecular Cell*, 9(2): 303 – 313.

Honeybee Genome Sequencing Consortium, 2006. Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature*, 443 (7 114): 931 – 949.

Horard B, Vanacker JM, 2003. Estrogen receptor-related receptors: Orphan receptors desperately seeking a ligand. *Journal of Molecular Endocrinology*, 31(3): 349 – 357.

Huss JM, Kopp RP, Kelly DP, 2002. Peroxisome proliferator-activated receptor coactivator-1alpha (PGC-1 alpha) coactivates the cardiac-enriched nuclear receptors estrogen-related receptor-alpha and -gamma. Identification of novel leucine-rich interaction motif within PGC-1alpha. *Journal of Biological Chemistry*, 277(43): 40 265 – 40 274.

Lu D, Kiriyama Y, Lee KY, Giguère V, 2001. Transcriptional regulation of the estrogen inclucible pS2 breast cancer marker gene by the ERR family of orphan nuclear receptors. *Cancer Research*, 61 (18): 6 755 – 6 761.

Luo J, Sladek R, Bader JA, Matthyssen A, Rossant J, Gigùere V, 1997. Placental abnormalities in mouse embryos lacking the orphan nuclear receptor ERR-beta. *Nature*, 388: 778 – 782.

Luo J, Sladek R, Carrier J, Bader JA, Richard D, Giguère V, 2003. Reduced fat mass in mice lacking orphan nuclear receptor estrogen-related receptor α. *Molecular and Cellular Biochemistry*, 23(22): 7 947 – 7 956.

Maglich JM, Sluder A, Guan X, Shi Y, McKee DD, Carrick K, Kamdar K, Willson TM, Moore JT, 2001. Comparison of complete nuclear receptor sets from the human, *Caenorhabditis elegans* and *Drosophila* genomes. *Genome Biology*, 2(8): 1 – 7.

Marchler-Bauer A, Bryant SH, 2004. CD-Search: Protein domain annotations on the fly. *Nucleic Acids Research*, 32: 327 – 331.

Mitsunaga K, Araki K, Mizusaki H, Morohashi K, Haruna K, Nakagata N, Giguere V, Yamamura K, Abe K, 2004. Loss of PGC-specific expression of the orphan nuclear receptor ERRβ results in reduction of germ cell number in mouse embryos. *Mechanics of Development*, 121(3): 237 – 246.

Ostberg T, Jacobsson M, Attersand A, Mata de Urquiza A, Jendeberg L, 2003. A triple mutant of the *Drosophila* ERR confers ligand-induced suppression of activity. *Biochemistry*, 42(21): 6 427 – 6 435.

Pettersson K, Svensson K, Mattsson R, Carlsson B, Ohlsson R, Berkenstam A, 1996. Expression of a novel member of estrogen response element-binding nuclear receptors is restricted to the early stages of chorion formation during mouse embryogenesis. *Mech. Dev.*, 54, 211 – 223.

Raghava GPS, 2000. APSSP: Advanced protein secondary structure prediction server. CASP4. 75 – 76.

Schäffer AA, Aravind L, Madden TL, Shavirin S, Spouge JL, Wolf YI, Koonin EV, Altschul SF, 2001. Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. *Nucleic Acids Research*, 29(14): 2 994 – 3 005.

Schreiber SN, Knutti D, Brogli K, Uhlmann T, Kralli A, 2003. The transcriptional coactivator PGC-1 regulates the expression and activity of the orphan nuclear receptor ERRα. *Journal of Biological Chemistry*, 278(11): 9 013 – 9 018.

Schwede T, Kopp J, Guex N, Peitsch MC, 2003. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research*, 31(13): 3 381 – 3 385.

（责任编辑：赵利辉）